3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

PERCEPTUAL PARSING OF NASAL VOWELS

John Ingram and Tom Mylne

Department of English
University of Queensland
Brisbane, Australia, 4072

ABSTRACT

An experiment is reported on the perception of nasal resonance by French listeners responding to gated monosyllabic words containing phonemically nasal or oral vowels, or vowels before nasal consonants. At issue, following Lahiri and Marslen-Wilson's [2] experiment on the perception of distinctive and allophonic nasality in Bengali and English, is the question of phonetic under-specificity in the recognition lexicon. However, we argue that the results are more informative of perceptual parsing strategies in the assignment of phonological structure to phonetic input, which argument is further supported by analysis of acoustic cues to oral-nasal resonance in the stimuli.

I. INTRODUCTION

Phonologists have long held the view that representations of items in the mental lexicon are phonetically under-specified, though the precise nature of this under-specificity has been the subject of much debate. On the other hand, Psycholinguistic and Phonetic investigations of word recognition have tended to the support the view that listeners exploit redundant phonetic features as soon as they become available in the stimulus, and this has been taken as prima facie evidence that items in the recognition lexicon are phonetically fully specified [3]. Thus, English listeners can detect an upcoming nasal consonant from nasal resonance in a preceding vowel, caused by a phonetic or phonological process of nasal spreading [1].

This property of utilising `non-distinctive' phonetic properties of speech, such as vowel nasal resonance in English, as soon as they become available to the auditory system, would seem to be dictated by considerations of information processing efficiency, and has been incorporated into the architecture of Marslen-Wilson's influential `cohort model' of lexical recognition, among others. If non-distinctive features are taken up by the perceptual mechanism, it seems natural to infer that redundant phonetic information is part of the phonetic constitution of items in the recognition lexicon, contrary to the tenet of classical and contemporary phonological theories that phonological representations are minimally redundant (phonetically under-specified).

However, Lahiri and Marslen-Wilson [2] (hereafter, L&MW) presented evidence on the perception of nasal vowels in Bengali, using a speech gating paradigm, which, they argued, supports phonological theory, that lexical items are phonetically under-specified, against the prevailing assumption in speech perception models, that items in the recognition lexicon are phonetically fully specified, or at least specified for some allophonic features. Bengali possesses distinctively nasal vowels, as well as having a phonological rule of nasal feature spreading. Thus, vowels in Bengali may be distinctively or allophonically nasalised. When a Bengali listener detects nasal resonance in a vowel, prior to the appearance of acoustic cues signalling the end of the vowel or the syllable coda, the stimulus is potentially ambiguous between a distinctively nasal vowel or one which is allophonically nasalised by an upcoming nasal consonant. This situation can be experimentally manipulated by the gating paradigm.

L&MW found that Bengali listeners interpreted nasal resonance as indicative of a (distinctive) nasal vowel, rather than as predictive of an upcoming nasal consonant, regardless of whether the [Cṽ..] stimulus had been gated from a word containing a nasal vowel (CṽC) or from one ending in a nasal consonant (CVN). L&MW argued that their results, which extended to Bengali and English listeners' perceptions of gated nasalised and non-nasalised vowels, strongly favoured a model of under-specified phonological representations (the UR model) over the so-called `surface representation' model (SR), in which redundant phonetic features, such as allophonic vowel nasalisation are specified in the recognition lexicon.

Recently, Ohala and Ohala [4] (hereafter, O&O) undertook a replication of the L&MW gating study, with certain methodological innovations, using Hindi and English listeners. Hindi, like Bengali, has distinctively nasal vowels and a rule of nasal spreading. Although their empirical findings substantially agreed with those of L&MW, O&O interpreted their results, on balance, as favouring the SR hypothesis, against the UR hypothesis.

It is notable that both of the studies

are perceptual in nature, without explicit attention to the acoustic cues mediating the perception of nasal vowels or vowel nasalisation. One possible source of differences between Bengali and Hindi listeners responses may reside in language specific differences in the implementation of distinctive and allophonic nasality. It was decided, therefore, to investigate the perception of nasal resonance in a language that has distinctive vowel nasality, but which is generally described as having little or no allophonic nasalisation, namely, French.

II. METHOD

2.1 Stimuli

Seven sets of French triplet stimuli were constructed, in which each member of a triplet shared the same initial consonant, and the same vowel (oral or nasal), but differed in the final consonant, which was either oral or nasal. An example set is the triplet *tâte* /tat/ (CVC), *tanne* /tan/ (CVN), *tante* /tãt/ (CṼC).

The procedure for constructing gated stimuli followed that of L&MW, but with variations to accommodate the durational characteristics of the French words. Gating points were established at zero crossings for every 4th glottal pulse from the onset of the vowel, using a waveform editor on a signal digitised at 20 kHz, with 12 bit resolution.

2.2 Subjects and Procedure:

The subjects for the listening experiment were 13 native French speakers, resident in Brisbane. All subjects were fluently bilingual, but used French regularly. Subjects were tested individually, hearing the stimuli over headphones. A response sheet was provided and subjects were asked to write down the word suggested to them by each stimulus, assuming that the stimulus was the beginning of a word ending in a consonant.

Experimental procedures also closely followed those of L&MW, contrasting in three potentially important respects, to innovations introduced by O&O. Following L&MW, but unlike the procedure adopted by O&O who used a forced choice paradigm, the response task was open-ended. Subjects were encouraged to guess the whole word on the basis of the truncated stimulus.

Also, as in L&MW's experiment, the stimuli from a gating set were presented sequentially, rather than randomised with items from other gating sets. Thus, listeners first heard the shortest gated stimulus (S1) in a set, made a judgement, and were then presented with the second shortest stimulus (S2), and so forth. Also in conformity with L&MW's experiment, the gated stimuli were simply truncated into silence, whereas O&O used amplitude modulated white noise to provide a more `natural' masking of the end of the syllable.

III. RESULTS

The key question of interest is how listeners differentially responded to the three types of gated vowels: oral vowels (derived from CVC words), distinctively nasal vowels (from CṼC words) and allophonically nasalised vowels (from CVN words, although nasalisation of such vowels in French is usually blocked). Responses to each gated stimulus were classified according to the type of word that was heard (CVC, CVN, or CṼC). Figure 1 shows the proportion of CVC, CVN and CṼC responses given to the three types of stimuli over successive gating intervals (gates 1-7).

For comparison with the Hindi and Bengali data, the percentage of response types to the three types of gated stimulus (CV(C), CṼ(C), and CV(N)), summed over successive gating intervals up to gate immediately preceding the vowel offset is shown in Table 1.

TABLE 1
Percentage of responses up to vowel offset

STIMULUS		RESPONSE TYPE		
		CVC	CṼC	CVN
CV(C)	Bengali	85	1	14
	Hindi	72	19	9
	French	93	4	3
CṼ(C)	Bengali	35	60	5
	Hindi	14	72	14
	French	49	48	3
CV(N)	Bengali	25	67	8
	Hindi	21	53	26
	French	86	8	6

Figure 1 shows that the gated CVC

stimuli, even from gate 1, yielded negligible occurrences of other than CVC responses, indicating that they were perceived overwhelmingly, as oral (non-nasal or non-nasalised) vowels.

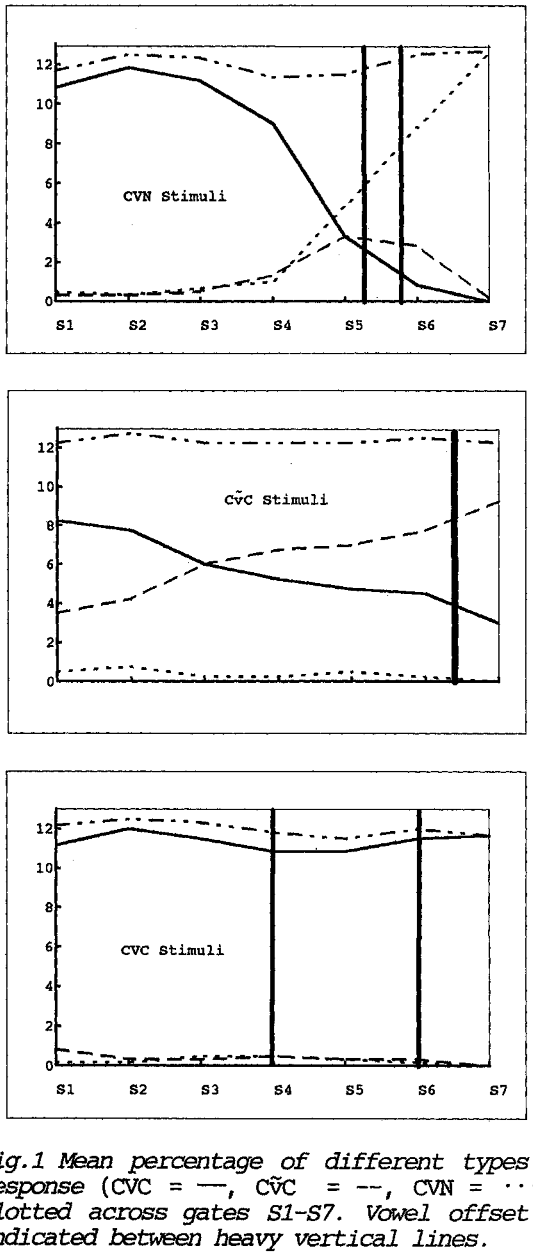

Fig.1 Mean percentage of different types of response (CVC = —, CṼC = --, CVN = ···) *plotted across gates S1-S7. Vowel offset is indicated between heavy vertical lines.*

Listeners' responses to gated French nasal vowels (CṼ(C)s) were similar to those obtained from Bengali and Hindi listeners, showing an increasing incidence of CṼC responses at the expense of CVC responses over successive gating intervals, while the proportion of CVN responses remained negligibly small. The graphical trend for the French CṼ(C) stimuli is virtually identical to that of L&MW's Bengali data.

Responses to French CV(N) stimuli were quite different from those of Bengali or Hindi, languages which have allophonic nasality, but were consistent with the expected effects of denasalisation (blocking of nasal spreading) in French. Rates of CVN or CṼC responses remained negligibly low across gating intervals for most stimuli in this class, until the end of the vowel was reached, at which time listeners detected unmistakable cues to the nasal consonant. Bengali and Hindi listeners, on the other hand, responded to nasal resonance in the gated vowel by treating it as distinctive vowel nasality, rather than the effects of nasal spreading.

IV. DISCUSSION

4.1 The underspecification hypothesis

Consider firstly, the responses to gated vowels where acoustic cues to oral-nasal resonance are in all likelihood absent, i.e.: CV(C) stimuli. In the Bengali data, L&MW found what they regarded as a significantly high proportion of CVN responses to these stimuli, roughly comparable with their estimated frequency of occurrence of CVN words in the Bengali lexicon. They reasoned, by the cohort model, that if the vowels in these items (CVN words) are unspecified for nasality in the recognition lexicon, they should remain active in the cohort until cues incompatible with consonantal nasal murmur are detected by the auditory system.

However, the findings from the Hindi and the French experiments do not sustain this prediction. In the case of the Hindi CV(C) targets, 19% of listeners responded CṼC and 9% as CVN. The French CV(C) data offer no support for the UR hypothesis either, which would predict that vowels in CVN words being unspecified for nasality in the recognition lexicon, should consequently remain in the cohort. However the rate of CVN responses to CV(C) stimuli was very low in the French data (4%).

The status of French CṼC words for the UR hypothesis is rendered problematical by competing phonological theories of French nasal vowels. By the `abstract generative' analysis CṼC words will be unspecified for nasality in lexical representations and should therefore remain in the cohort. By the `concrete' phonological analysis, they will be specified for nasality and, presumably be excluded from the cohort on the UR model.

Turning to those stimuli in which listeners can reasonably be assumed to detect oral-nasal resonance over the up-to-offset gating sequence, (namely the Bengali, Hindi, and French CṼ(C)s, and the Bengali and Hindi, but not the French, CV(N)s), the data indicate that when confronted with nasal resonance which is ambiguous between a distinctive nasal vowel and an upcoming nasal consonant, listeners will opt for the former, until or unless subsequent phonetic cues force abandonment of the hypothesis.

4.2 Parsing strategies

Perceptual processing of the speech signal inevitably involves abstracting away from the phonetic detail to the distinctive properties that differentiate items in the listener's lexicon. Word recognition is therefore inevitably guided by a phonological representation which is abstract, and in this respect, our account of the data is not incompatible with L&MW's UR hypothesis.

The perceptual data are consistent with a simple heuristic or parsing strategy, a phonological `minimal distance principle', that, by default, assigns a phonetic feature detected in the speech input stream to the phonological identity of the current acoustic segment. This parsing strategy needs to be invoked in languages like Bengali or Hindi or which have both distinctive and allophonic vowel nasalisation. Vowel nasal resonance is unambiguously distinctive in French and allophonic in English.

4.3 Denasalisation in French CVNs

A characteristic spectral peak at approximately 250 Hz in gated CṼCs provided the basis for an acoustic measure of vowel nasal resonance. A three way analysis of variance was undertaken with stimulus type (CṼC, CVN and CVC), vowel type (/a/, /e/, and /ɔ/), and gating points (gates 1-3) as independent variables. Two of the three vowels showed clear denasalisation (/a/ and /ɔ/). The one that did not (/e/) also showed an elevated incidence of CṼC responses to CV N stimuli. We take this as further, though provisional, support for the proposed parsing strategy.

REFERENCES

[1] Ali, H.L., Gallagher, T., Goldstein, J., and Daniloff, R.G. Perception of coarticulated nasality. *J. Acous. Soc. Am.* 49: 538-540, 1971.

[2] Lahiri, A., and Marslen-Wilson, W. D. The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38: 245-294, 1991.

[3] Ohala, J. J. Comments on Chapter 9. In G. J. Docherty and D. R. Ladd, Papers in Laboratory Phonology II: Gesture, Segment, Prosody (pp. 255-257). Cambridge: Cambridge University Press, 1992.

[4] Ohala, J. J. and Ohala, M. (1993). Speech perception and lexical representation: the role of vowel nasalisation in Hindi and English. 2nd Draft of Paper presented at 4th Conference on Laboratory Phonology, Oxford, August 1993.